تحلیل آماری

نام استاد: بختیاری

گردآورنده: فرخ سیفی

۱- **نام کتاب:** آمار و کاربرد آن در مدیریت

نویسنده : دکتر عادل آذر و دکتر منصور مومنی

ترجمه : ....

انتشارات : سمت

مبانی تئوری تخمین و آزمون فرض

تخمین : سوال

فرض : فرضیه

تئوری تخمین بدنبال پاسخگویی به سوالات آماری است .

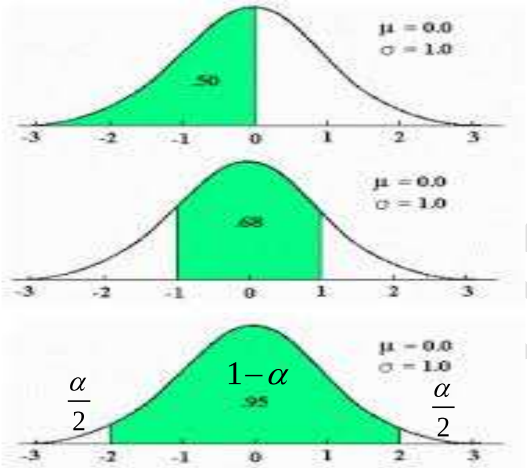
تخمین فاصله ای میانگین جامعه (μ_x)

	جامعه	نمونه
میانگین	μ_x	$\bar{x}$
میانه	md	md
واریانس	σ_x^2	s_x^2

استنباط آماری :

تعمیم آماره های بدست آمده از نمونه جامعه

در تخمین فاصله ای میانگین جامعه بدنبال ای هستیم که بر حسب میانگین نمونه ($\bar{x}$) و با درجه اطمینان مناسب میانگین جامعه را تخمین می زنیم این تخمین میبایستی بصورت فاصله و نه نقطه ای باشد بعبارت دیگر میخواهیم بگوئیم $\bar{x}-\in\leq\mu_x\leq\bar{x}+\in$ باشد .

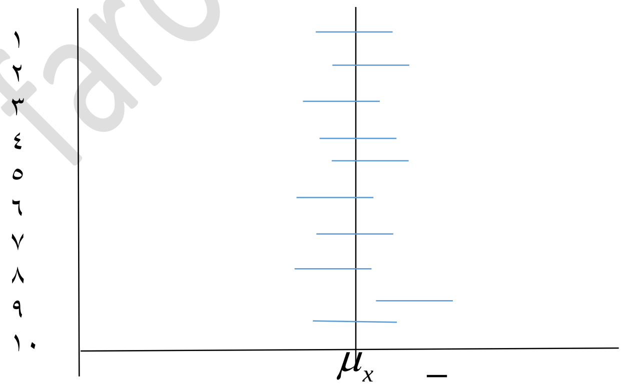


خطا α

α = ۰/۰۵

۱- α =۰,۹۵

فرض کنید محققی به تخمین میانگین سود هفتگی شرکت تمایل دارد اگر نمونه شامل ۲۰ مشاهده هفتگی باشد و عمل نمونه گیری ۱۰ بار تکرار شده باشد و در هر بار تخمین فاصله میانگین جامعه μ_x بدست آمده باشد در قالب شکل زیر تفسیر میشود .



نکته :

تخمین فاصله ای μ_x یا مقدار ε (اپسیلن) تحت تاثیر سطح اطمینان و توزیع $\bar{x}$ است .

توزیع $\bar{x}$ با شرایط زیر تعیین میشود .

۱- نوع توزیع جامعه آماری که نرمال یا غیر نرمال باشد .

۲- کیفیت انحراف معیار جامعه که معلوم یا نامعلوم باشد .

۳- اندازه نمونه که کوچک یا بزرگ باشد .

حد مرکزی :

هر گاه تعداد نمونه بزرگ باشد فارغ از توزیع جامعه میانگین نمونه دارای توزیع نرمال خواهد بود و داریم :

$$\mu_{\bar{x}} = \mu_x$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

<u>۱-۱ - توزیع جامعه آماری نرمال با انحراف معیار معلوم</u>

$$\pm z_{\frac{\alpha}{2}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$or$$

$$z = \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}}$$

$$p(-z_{\frac{\alpha}{2}} \le z \le z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$p(-z_{\frac{\alpha}{2}} \le \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}} \le z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$p(-z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}} - \bar{x} \le -\mu_x \le -\bar{x} + z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}}) = 1 - \alpha$$

$$p(-\bar{x} - z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}} + \ge -\mu_x \ge -\bar{x} + z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}}) = 1 - \alpha$$

$$p(\bar{x} - z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}} \le \mu_x \le \bar{x} + z_{\frac{\alpha}{2}} . \frac{\sigma_x}{\sqrt{n}} = 1 - \alpha$$

مثال :

بررسیها نشان می دهد توزیع وزن تولید شده در یک کارخانه بزرگ نرمال و انحراف معیار آن ۲۱ تن باشد از آنجا که اندازه گیری وزن محصولات بطور روزانه امکان پذیر نیست یک نمونه ۵۰ روزه از تولیدات انتخاب شده است که میانگین وزن آن ۸۷۱ تن است در سطح اطمینان ۹۰٪ است مطلوبست میانگین واقعی وزن محصولات تولید شده در یک روز

$$\bar{x} = 871$$
$$n = 50$$
$$\sigma_x = 21$$
$$1-\alpha = 90\%$$
$$\alpha = 10\%$$
$$\pm z_{\frac{\alpha}{2}} \Rightarrow z_{0.05} = \pm 1.645 \rightarrow optainindiagram(z)$$
$$p(871 - 1.645 \times \frac{21}{\sqrt{50}} \le \mu_x \le 871 + 1.645 \times \frac{21}{\sqrt{50}}$$
$$866.11 \le \mu_x \le 875.89$$

۲-۱- توزیع جامعه آماری نرمال با انحراف معیار نامعلوم

هرگاه انحراف معیار جامعه ای نرمال نامعوم باشد در تخمین فاصله ای μ_x بجای σ_x^2 از $s_{\bar{x}}^2$ استفاده میکنیم این عمل موجب خواهد شد بجای رابطه $\frac{\bar{x}-\mu_{\bar{x}}}{s_{\bar{x}}} \rightarrow \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$ استفاده کنیم در این حالت نیز توزیع $\bar{x}$ توزیع نرمال میباشد که بواسطه نامعلوم بودن σ_x پراکندگی بیشتری نسبت به توزیع Z دارد .

df= درجه آزادی =n-۱

$$p(-t_{\frac{\alpha}{2}}, df \le t \le t_{\frac{\alpha}{2}}, df) = 1-\alpha$$

اگر تعداد نمونه کمتر از ۳۰ تا بود

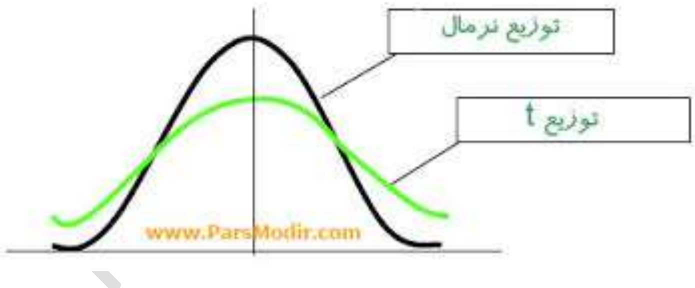
$$p(\bar{x} - t_{\frac{\alpha}{2},df} . s_{\bar{x}} \le \mu_x \le \bar{x} + t_{\frac{\alpha}{2},df} . s_{\bar{x}}) = 1-\alpha$$

اگر تعداد نمونه بیشتر از ۳۰ تا بود

$$p(\bar{x} - z_{\frac{\alpha}{2}} . s_{\bar{x}} \le \mu_x \le \bar{x} + z_{\frac{\alpha}{2}} . s_{\bar{x}}) = 1-\alpha$$

نکته :

بر اساس قضیه حد مرکزی چنانچه حجم نمونه بزرگ باشد توزی t استیودنت به سمت توزیع نرمال میل میکند لذا میتوان بجای توزیع t از z برای تخمین μ_x استفاده کرد .



مثال :

بازاریابی درصدد برآورد قدرت خرید ساکنان یک محله در تهران است او ناچار باید یک نمونه تصادفی ۱۰ تایی از بین خریداران انتخاب و قدرت خرید هر یک را اندازه گیری کند قدرت خرید نمونه فوق بر حسب ۱۰ هزار تومان چنین است .

Xi=۸،۷،۵،۴،۱۲،۱۵،۱۰،۱۳،۱۴،۱۲

قدرت خرید ساکنین محله از توزیع نرمال برخوردار است درسطح اطمینان ۹۵٪ میانگین قدرت خرید آنها را برآورد کنید .

N=۱۰ α =۰,۰۵   ۱-α =۰,۹۵

$$\overline{P}=(\overline{x}-t_{\frac{\alpha}{2}},d.f.s_{\bar{x}}\le \mu_x\le \overline{x}+t_{\frac{\alpha}{2}},d.f.s_{\bar{x}})=1-\alpha$$

$$\bar{x}=\frac{\sum x_i}{10}=\frac{100}{10}=10$$

$$sx=\sqrt{\frac{\sum (x_i-\bar{x})^2}{n-1}}=\sqrt{\frac{(8-10)^2+(7-10)^2+(8-5)^2+(8-4)^2,....}{9}}=\sqrt{\frac{132}{9}}=3.83$$

$$d_f=n-1=10-1=9$$

$$s_{\bar{x}}=\frac{sx}{\sqrt{n}}=\frac{3.83}{\sqrt{10}}=1.211$$

$$t_{\frac{\alpha}{2}},df=t_{0.025},9=\pm 2.262\rightarrow obtainofdiagram$$

$$p(10-2.262\times 1.211\le \mu_x\le 10+2.262\times 1.211=0.95$$

$$7.261\le \mu_x\le 12.739$$

توزیع جامعه غیر نرمال :

در این حالت چنانچه اندازه نمونه (N) بزرگتر از ۳۰ باشد ($n\ge 30$) در صورتی که انحراف معیار جامعه معلوم باشد از آماره z و در صورتی که نامعلوم باشد از آماره t استفاده میکنیم .

$$n\ge 30 \begin{cases} \sigma_x \text{ معلوم } \longrightarrow z \\ \sigma_x \text{ نامعلوم } \longrightarrow t \end{cases}$$

چنانچه حجم نمونه کوچک باشد $n\le 30$ برای تخمین میانگین جامعه نمیتوانیم از z یا t استفاده کنیم در ای حالت قضیه چی پی شف در تخمین میانگین جامعه بکار میرود .

قضیه چی بی شف :

احتمال قرار گرفتن میانگین نمونه در بین k انحراف معیار استاندارد $\sigma_{\bar{x}}$ برابر است با احتمال اینکه :

$$p(\left[\overline{x}-\mu_x\right]\le k.\sigma\overline{x})\ge 1-\frac{1}{k^2}or1-\alpha$$

$$p(\overline{x}-k.\sigma\overline{x}\le \mu_x\le \overline{x}+k.\sigma\overline{x}\ge 1-\alpha$$

اگر انحراف معیار جامعه نامعلوم باشد از رابطه زیر استفاده میکنیم

$$p(\overline{x}-k.s\overline{x}\le \mu_x\le \overline{x}+k.s\overline{x}\ge 1-\alpha$$

طبق قضیه چی پی شف فاصله میانگین نمونه از میانگین جامعه بر اساس انحراف معیار قابل تخمین است بعبارت دیگر احتمال فاصله $\overline{x}$ از μ_x برابر $(\left|\overline{x}-\mu_x\right|)$ k انحراف معیار سمت چپ و k انحراف معیار راست قرار گیرد حداقل $1-\frac{1}{k^2}$ میباشد ، با برابر قرار دادن $1-\alpha$ با $1-\frac{1}{k^2}$ میتوانیم به رابطه ای جهت تخمین جامعه برسیم .

مثال :

از شیشه هایی که با یک دستگاه پر شده است یک نمونه تصادفی ۲۵ تایی انتخای میکنیم میانگین آن ۲۵۰ میلی لیتر و انحراف معیار آن ۱۰ میلی لیتر است هیچ دلیلی بر نرمال بودن توزیع مایع ریخته شده در داخل شیشه ها وجود ندارد در سطح اطمینان ۹۵٪ میانگین کل مایع ریخته شده در شیشه ها را برآورد کنید .

$$1-\frac{1}{k^2}=1-\alpha \rightarrow \cancel{1}-\frac{1}{k^2}=\cancel{1}-0.05$$

$$k^2=\frac{1}{0.05} \Rightarrow k=\sqrt{\frac{1}{0.05}}=\pm 4.47$$

$$p(250-4.47\times\frac{10}{\sqrt{25}} \le \mu_x \le 250+4.47\times\frac{10}{\sqrt{25}} \ge 0.95$$

$$241.6 \le \mu_x \le 258.94$$

تخمین فاصله تفاضل میانگین دو جامعه :

برای مقایسه میانگین دو جامعه میتوان بصورت جداگانه میانگین ها را طبق مباحثی که قبلاً گفته شد تخمین زد و نتایج را با هم مقایسه نمود بجای دوبار تخمین زدن بهتر است از تخمین تفاضل میانگین دو جامعه استفادن نمائیم اگر $\overline{x}_1, \overline{x}_2$ دو متغییر تصادفی مستقل از هم باشند داریم .

$$\mu_{\overline{x}_1-\overline{x}_2}=\mu_{\overline{x}_1}-\mu_{\overline{x}_2}$$

$$\mu_{\overline{x}_1-\overline{x}_2}=\mu_1-\mu_2$$

$$\sigma^2_{\overline{x}_1-\overline{x}_2}=\sigma^2_{x_1}+\sigma^2_{x_2}$$

$$\sigma_{\overline{x}_1-\overline{x}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

توزیع دو جامعه آماری نرمال و انحراف معیار (σ_1, σ_2) معلوم باشد .

$$p\left(-z_{\frac{\alpha}{2}} \le z \le z_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$p\left(-z_{\frac{\alpha}{2}} \le \left(\overline{x}_1-\overline{x}_2\right)-\left(\mu_1-\mu_2\right) \le z_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$p\left(\left(\overline{x}_1-\overline{x}_2\right)-z_{\frac{\alpha}{2}}.\sigma_{\overline{x}_1-\overline{x}_2} \le \left(\mu_1-\mu_2\right) \le \left(\overline{x}_1-\overline{x}_2\right)+z_{\frac{\alpha}{2}}.\sigma_{\overline{x}_1-\overline{x}_2}\right)=1-\alpha$$

مثال :

هدف از تحقیقی مقایسه عملکرد کارمندان در دو سازمان است و ملاک مقایسه میانگین دو جامعه آماری است عملکرد کلیه کارمندان در دسترس نبوده و ناچار مجبور به نمونه گیری از هر جامعه هستیم از سازمان الف یک نمونه تصادفی ۲۵ نفره انتخاب شده که میانگین آن ۶۰ میباشد و از سازمان ب یک نمونه تصادفی ۲۰ نفره انتخاب شده که میانگین آن ۵۵ است توزیع نمره های عملکرد هر دو سازمان نرمال و انحراف معیار عملکرد به ترتیب ۱۰ و ۱۲ است در سطح اطمینان ۹۹٪ میانگین عملکرد هر دو جامعه را با هم مقایسه کنید .

سازمان الف	سازمان ب
$n_1=25$	$n_2=20$
$\overline{x}_1=60$	$\overline{x}_2=55$
$\sigma_1=10$	$\sigma_2=12$
$normal$	$normal$

$$1-\alpha=99\% \rightarrow \alpha=0.01$$

$$z_{\frac{\alpha}{2}}=0.005 \Rightarrow \pm 2.58$$

$$\sigma \overline{x}_1-\overline{x}_2=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=\sqrt{\frac{10^2}{25}+\frac{12^2}{20}}=\sqrt{4+7.2}=\sqrt{11.2}=3.35$$

$$P(60-55)-2.58\times 3.35 \le \mu_1-\mu_2 \le (60-55)+2.58\times 3.35=0.99$$

$$-3.634 \le \mu_1-\mu_2 \le 13.634$$

تحلیل : بدلیل وجود یک منفی و یک مثبت تفاوت معنی داری وجود ندارد.

نکته :

در تخمین فاصله ای $\mu_1-\mu_2$ در صورتی که هر دو دامنه مثبت باشد یعنی هم حد بالا و هم حد پائین مثبت باشند میتوانیم با سطح اطمینان مورد نظر بگوئیم که $\mu_1 \succ \mu_2$ میباشد .

در صورتی که هر دو دامنه منفی باشند میتوانیم با سطح اطمینان مورد نظر بگوئیم که $\mu_1 \prec \mu_2$ میباشد .

در حالتی که حد پائین منفی و حد بالا مثبت باشد چون تخمین بعمل آمده از نقطه صفر میگذرد بین μ_1,μ_2 تفاوت معنی داری وجود ندارد .

توزیع دو جامعه آماری نرمال و انحراف معیار نامعلوم

$$t=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{s_{\overline{x}_1-\overline{x}_2}}$$

$$s_{\overline{x}_1-\overline{x}_2} \rightarrow \sigma_1^2=\sigma_2^2$$

$$s_{\overline{x}_1-\overline{x}_2} \rightarrow \sigma_1^2 \ne \sigma_2^2$$

$$if\,\sigma_1^2=\sigma_2^2 \Rightarrow s_{\overline{x}_1-\overline{x}_2}=sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$s^2p=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} \rightarrow sp=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$

$$if\,\sigma_1^2 \ne \sigma_2^2 \Rightarrow s_{\overline{x}_1-\overline{x}_2}=sp\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$\sigma_1^2=\sigma_2^2 \Rightarrow P\left((x_1-x_2)+t_{\frac{\alpha}{2}},d.f.sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1-\mu_2 \le (x_1-x_2)-t_{\frac{\alpha}{2}},d.f.sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

$$\sigma_1^2 \ne \sigma_2^2 \Rightarrow P\left((x_1-x_2)+t'_{\frac{\alpha}{2},\mathrm{d.f}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le \mu_1-\mu_2 \le (x_1-x_2)-t'_{\frac{\alpha}{2},\mathrm{d.f}}.\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)=1-\alpha$$

$$d.f'=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_1}\right)^2}{n_2-1}}$$

مثال :

هدف از تحقیقی مقایسه سطح آمادگی کارمندان در سازمان الف با سازمان ب است در این تحقیق از سازمان الف یک نمونه ۹ نفره انتخاب شده که میانگین آن ۴۵ و انحراف معیارش ۱۲ است در حالی که میانگین و انحراف معیار سطح آمادگی کارمندان سازمان ب در یک نمونه ۱۵ نفره به ترتیب ۵۵ و ۱۴ است فرض کنید توزیع نمره های سطح آمادگی نرمال و $\sigma_1^2=\sigma_2^2$ است در سطح اطمینان ۹۰٪ تخمین لازم را برای مقایسه میانگین دو جامعه بعمل آورید .

$$SP=\sqrt{\frac{8(12)^2+14(14)^2}{9+15-2}}=13.31$$

$$\frac{\alpha}{2}=0.05$$

$$d.f=n_1+n_2-2=9+15-2=22$$

$$T0.05,22=\pm1.717$$

$$(45-55)\pm1.717(13.31\times\sqrt{\frac{1}{9}+\frac{1}{15}})=(-19.635-0.634)$$

$$-19.635\le\mu_1-\mu_2\le-0.364$$

تحلیل : چون هر دو دامنه منفی است ، پس در سطح خطای ۱۰٪ میتوان گفت که $\mu_1\prec\mu_2$ است .

تخمین فاصله ای نسبت موفقیت جامعه (p) :

بسیاری از تحقیقات بدنبال تخمین نسبت عناصر از جامعه آماری هستند که دارای یک ویژگی مورد نظر میباشند همچنین بسیاری از تحقیقات از مقیاس اسمی یا رتبه ای برخوردارند که پارمتر توصیف آنها نسبت موفقیت است مواردی همچون درصد کالاهای معیوب ، درصد کارکنان که از کار خود راضی اند و درصد مدیران که وظیفه مدارند اینها مصادیق تحقیقاتند .

برای تخمین پارامتر $p=\frac{x}{n}$ از آماره $\overline{p}$ استفاده میکنیم توزیع نمونه گیری $\overline{p}$ در نمونه های بزرگ از تقریب نرمال برخوردار است و متغییر استاندارد ما چنین است .

$$z=\frac{\overline{p}-p}{\sqrt{\frac{\overline{p}(1-p)}{n}}}$$

$$z=\frac{\overline{p}-\mu_{\overline{p}}}{\sigma_{\overline{p}}}$$

$$\varepsilon(\overline{P})=\mu_{\overline{P}}=P$$

$$\sigma_{\overline{P}}=\sqrt{\frac{\overline{P}(1-P)}{n}}$$

$$z=\frac{\overline{p}-p}{\sqrt{\frac{\overline{p}(1-p)}{n}}}$$

$$\overline{p}\left(\overline{p}-z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{P}(1-P)}{n}}\le p\le\overline{p}+z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{P}(1-P)}{n}}\right)=1-\alpha$$

مثال :

هدف از یک تحقیق تعیین نسبت افراد ناراضی در سازمان است از آنجا که دسترسی به افراد سازمان میسر نیست محققان ۴۰۰ نفر از کارمندان را بطور تصادفی انتخاب کرده اند که فقط ۳۲ نفر از آنها از کار خود ناراضی اند نسبت افراد ناراضی را در سازمان در سطح خطای یک درصد برآورد کنید .

$$n = 400$$
$$x = 32$$
$$\overline{p} = \frac{x}{n} = \frac{32}{400} = 0.08$$
$$1-\alpha = 0.99$$
$$\alpha = 0.01$$
$$\frac{\alpha}{2} = 0.005$$
$$z_{0.005} = \pm 2.58$$
$$p\left(0.08 - 0.035 \le p \le 0.08 + 0.035\right) = 0.99$$
$$0.045 \le p \le 0.115$$

تخمین فاصله تفاضل موفقیت دو جامعه :

$$p\left[\left(\overline{p}_1 - \overline{p}_2\right) - z_{\frac{\alpha}{2}} . s_{\overline{p}_1 - \overline{p}_2} \le p_1 - p_2 \le \left(\overline{p}_1 - \overline{p}_2\right) + z_{\frac{\alpha}{2}} . s_{\overline{p}_1 - \overline{p}_2}\right] = 1 - \alpha$$
$$s_{\overline{p}_1 - \overline{p}_2} = \sqrt{\frac{\overline{p}_1\left(1-\overline{p}_1\right)}{n_1} + \frac{\overline{p}_2\left(1-\overline{p}_2\right)}{n_2}}$$

مثال :

در تحقیقی مقایسه درصد مدیران برخوردار از روش S۱ در سازمانهای دولتی با مدیران سازمانهای خصوصی برنامه ریزی شده است بدین منظور از میان مدیران دولتی یک نمونه ۵۰۰ تایی و از مدیران خصوصی یک نمونه ۴۰۰ تایی انتخاب شده است ۱۰۰ نفر از مدیران دولتی و ۵۰ نفر از مدیران خصوصی دارای روش S۱ هستند با در نظر گرفتن سطح اطمینان ۹۵٪ نسبت واقعی مدیران برخوردار از روش S۱ را در دو جامعه آماری مقایسه کنید .

$$n_1 = 500 \qquad n_2 = 400$$
$$x_1 = 100 \qquad x_2 = 50$$
$$\overline{p}_1 = \frac{100}{500} = 0.2 \qquad \overline{p}_2 = \frac{50}{400} = 0.125$$
$$z_{0.025} = \pm 1.96$$
$$\left(0.2 - 0.125\right) \pm 1.96\sqrt{\frac{0.2(0.8)}{500} + \frac{0.125(0.875)}{400}}$$
$$0.027 \le p_1 - p_2 \le 0.123$$

در سطح اطمینان ۹۵٪ نسبت واقعی مدیران برخوردار از روش S۱ در سازمانهای دولتی بیش از سازمانهای خصوصی است زیرا هر دو دامنه برآورد تفاضل $p_1 - p_2$ اعداد مثبت هستند .

تخمین فاصله ای واریانس جامعه σ_x^2

$$\chi^2=\frac{(n-1)s_x^2}{\sigma_x^2}$$

$$p\left(\chi^2_{1-\frac{\alpha}{2},d.f}\le\chi^2\le\chi^2_{\frac{\alpha}{2},d.f}\right)=1-\alpha$$

$$p\left(\frac{(n-1)s_x^2}{\chi^2_{\frac{\alpha}{2},d.f}}\le\sigma_x^2\le\frac{(n-1)s_x^2}{\chi^2_{1-\frac{\alpha}{2},d.f}}\right)=1-\alpha$$

مثال :

یکی از کارگزاران بازار بورس تهران درصدد تعیین ریسک شرکتهایی است که سهام خود را عرضه میکنند در این زمینه یگ نمونه تصادفی ۱۵ تایی از بین شرکتها انتخاب کرده که میانگین و واریانس سود سالانه آنها به ترتیب ۲۵۰۰۰ و ۱۲۲۵ میباشد توزیه سود سالانه شرکتها از تقریب نرمال برخوردار است در سطح خطای ۵٪ حدود ریسک شرکتها را بدست آورید .

$$\alpha=0.05$$

$$\frac{\alpha}{2}=0.025$$

$$\chi^2_{0.025,14}=26.119$$

$$\chi^2_{0.975,14}=5.629$$

$$p\,\frac{(14\times1225)}{26.119}\le\sigma_x^2\le\frac{(14\times1225)}{5.629}\Rightarrow 656.61\le\sigma_x^2\le3046.72$$

$$jazr.\text{gerefte}.\text{shode}\Rightarrow 25.62\le\sigma_x\le55.20$$

تخمین فاصله ای نسبت واریانس جامعه آماری $\frac{\sigma_1^2}{\sigma_2^2}$

$$F=\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_1^2}}=\frac{\sigma_2^2s_1^2}{\sigma_1^2s_2^2}$$

$$p\left(\frac{\frac{s_1^2}{s_2^2}}{F_{\frac{\alpha}{2},d.f_1,d.f_2}}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{s_1^2}{s_2^2}.F_{\frac{\alpha}{2},d.f_2,d.f_1}\right)=1-\alpha$$

مثال :

هدف از تحقیقی مقایسه پراکندگی نمره های دانشجویان در دو دانشکده الف و ب است در این تحقیق از دانشکده الف یک نمونه تصادفی ۱۶ تایی انتخاب شده که میانگین آن ۱۴ و واریانس آن ۱۶ است در دانشکده ب نیز نمونه ۱۰ تایی با میانگین ۱۲ و ۱۵ انتخاب شده است در سطح اطمینان ۹۰٪ پراکندگی نمونه ها را در دو دانشکده مقایسه کنید .

$$d.f_1 = n_1 - 1 \Rightarrow 15 - 1 = 14$$
$$d.f_2 = n_2 - 1 \Rightarrow 10 - 1 = 9$$
$$\alpha = 0.10$$
$$\frac{\alpha}{2} = 0.05$$
$$F_{0.05, d.f_1} = F_{0.05,15\times9} = 3.01$$
$$F_{0.05,9\times15} = 2.59$$
$$p\left(\frac{\frac{16}{12}}{3.01} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{16}{12}.2.59\right) = 90\% \Rightarrow p\left(0.433 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.453\right) = 90\%$$

نکته :

در تحلیل فاصله اطمینان نسبت واریانس دو جامعه یکی از موارد سه گانه زیر امکانپذیر است .

۱- اگر هر دو دامنه بزرگتر از ۱ باشند در سطح اطمینان مورد نظر میتوان گفت $\sigma_1^2 \succ \sigma_2^2$ میباشد.

۲- اگر هر دو دامنه کوچکتر از ۱ باشند در سطح اطمینان مورد نظر میتوان گفت $\sigma_1^2 \prec \sigma_2^2$ میباشد.

۳- در حالتی غیر از این موارد یک و دو میتوان گفت تفاوت معنی داری بین تراکنشهای دو جامعه وجود ندارد .

- در مثال بالا تفاوت معنی داری وجود ندارد .

آزمون فرض آماری :

فرمول آمار استباطی برای پاسخ به سوالات تحقیق همان فرمول تخمین آماری است در حالی که اگر از مرحله سوال عبور کرده و فرضیه ای تعریف آن ((حدس عالمانه و هوشمندانه)) داشته باشیم از فنون آزمون فرض آماری برای تحلیل مسائل استفاده میکنیم و مراحل اجرای آزمون بصورت زیر است ؛

۱- فرضیات پژوهشی را به فرضیات آماری تبدیل نمائید (H_0, H_1)

۲- تعیین توزیع نمونه گیری آماره و نوع آماره آزمون

۳- تعیین سطح زیر منحنی (H_0, H_1) و محاسبه مقدار بحرانی

۴- مقایسه آماره آزمون و مقدار بحرانی و تصمیم گیری درخصوص تائید یا رد H_0

نکته :

قاعده پذیرفته شده در آمار اینست که H_0 آزمون میشود و بر اساس تائید یا رد H_0 فرضیه پژوهشی تحلیل میشود آنچه که آزمون میشود باید دارای تساوی باشد بنابراین پس از تدوین فرضیات آماری آن یک را که دارای مساوی باشد H_0 مینامیم میتواند ادعا یا نقیض ادعا باشد بعنوان مثال در فرضیه پژوهشی نسبت مدیران برخوردار از روش S1 در کشور بیش از ۶۰٪ است داریم ؛

$p > 60\% \rightarrow H_1$ ادعا

$p \le 60\% \rightarrow H_0$ نقیض ادعا

مثال :

فرضیه پژوهشی بدین صورت تدوین یافته است میانگین نمره مسئولیت پذیری مدیران سازمان حداقل ۵۰ است .

$H_0 \rightarrow \mu_x \ge 50$ ادعا

$H_1 \rightarrow \mu_x < 50$ نقیض ادعا

خطاهای آماری :

گزینه های صحیح		نتیجه گیری از نمونه
H_0 غلط است	H_0 درست است	
خطای نوع دوم (β)	تصمیم درست $(1-\alpha)$	H_0 پذیرفته میشود
تصمیم درست توان آزمون $(1-\beta)$	خطای نوع اول (α)	H_0 رد میشود

توزیع نمونه گیری آماره :

پارامتر فرضیه پژوهش	نماد آماری فرضیه	آماره	توزیع نمونه گیری	آماره آزمون
میانگین جامعه	μ_0	$\bar{x}$	t,z	t,z
مقایسه میانگین دو جامعه	$\mu_1-\mu_2$	$\bar{x}_1-\bar{x}_2$	t,z	t,z
zنسبت موفقیت جامعه	p_0	$\bar{p}$	نرمال	z
نسبت موفقیت دو جامعه	p_1-p_2	$\bar{p}_1-\bar{p}_2$	نرمال	z
واریانس جامعه	σ^2	s_x^2	کای مربع	χ^2
مقایسه واریانس دو جامعه	$\frac{\sigma_1^2}{\sigma_2^2}$	$\frac{s_1^2}{s_2^2}$	فیشر	F

آزمون فرض یک دنباله و دو دنباله

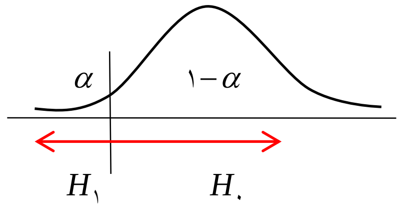
در آزمون فرض آماری H_0 را در سطح اطمینان در نظر میگیریم و H_1 را سطح خطا بنابراین همیشه H_1 که سطح خطا بوده و مایل به رد آن هستیم در انتهای دنباله های توزیع قرار میگرد و بسته به نوع آزمون یک دنباله یا دو دنباله میتواند باشد که در قالب مثالهای زیر توضیح داده میشود .

مثال :

فرضیه پژوهش بدین صورت تدوین یافته است (میانگین دستمزد یک کارخانه صنعتی کمتر از ۲۰۰۰۰ تومان است) .

آزمون یک دنباله سمت چپ

$$H_1 \to \mu_x < 20000$$
$$H_0 \to \mu_x \geq 20000$$

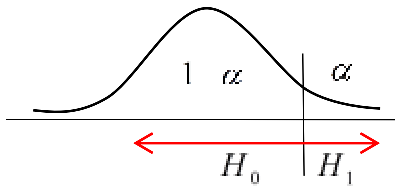


مثال :

فرضیه پژوهش بدین صورت تدوین یافته است (حداکثر ۱۰٪ محصولات کارخانه الف معیوب هستند)

آزمون یک دنباله سمت راست

$$H_0 \to p \leq 0.10$$
$$H_1 \to p > 0.10$$

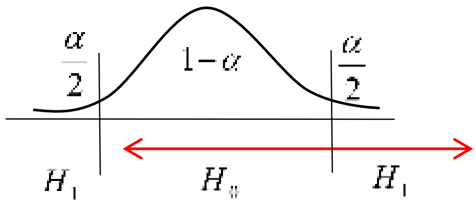


مثال :

میانگین نمره مسئولیت پذیری کارمندان مرد با نمره مسئولیت پذیری کارمندان زن تفاوتی ندارد .

آزمون دو دنباله

$$H_0 \rightarrow \mu_1 = \mu_2$$
$$H_1 \rightarrow \mu_1 \neq \mu_2$$



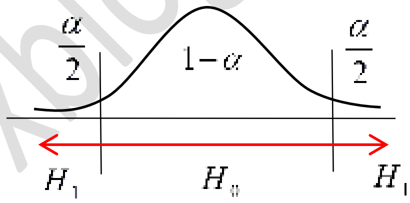
آزمون فرض آماری میانگین یک جامعه μ_x

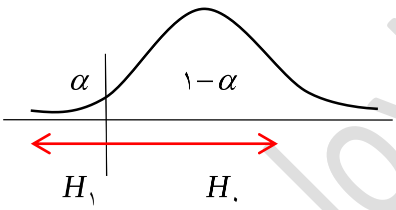
دراین آزمون سه حالت بوجود می آید

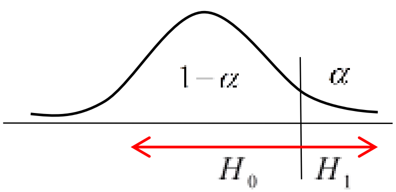
$$H_0 \rightarrow \mu_x = \mu_0$$
$$H_1 \rightarrow \mu_x \neq \mu_0$$

$$H_0 \rightarrow \mu_x \geq \mu_0$$
$$H_1 \rightarrow \mu_x < \mu_0$$

$$H_0 \rightarrow \mu_x \leq \mu_0$$
$$H_1 \rightarrow \mu_x > \mu_0$$







جامعه نرمال و انحراف معیار معلوم

$$z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$$

جامعه نرمال و انحراف معیار معلوم $n \leq 30$

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$$

جامعه نرمال و انحراف معیار نامعلوم $n > 30$

$$z = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$$

مثال :

فرضیه ای بدین صورت تدوین یافته ، میانگین مسئولیت پذیری مدیران در شرکت حداقل ۵۰ است محقق برای بررسی فرضیه یک نمونه ۶۴ تایی از بین مدیران کشور انتخاب کرده که میانگین و انحراف معیار به ترتیب ۴۵ و ۱۶ است ، در سطح خطای ۵٪ صحت قضیه فوق را بررسی کنید .

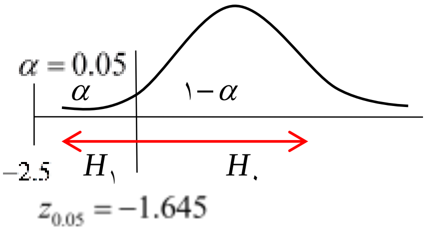
۱- فرضها

$$H_0 \rightarrow \mu_x \geq 50$$
$$H_1 \rightarrow \mu_x < 50$$

۲- آماره آزمون

$$z = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{45-50}{\frac{16}{\sqrt{64}}} = \frac{-5}{2} = -2.5$$

۳- محاسبه مقدار بحرانی



۴- تصمیم گیری

چون -2.5 در ناحیه H_1 قرار گرفته پس ادعا رد میشود .

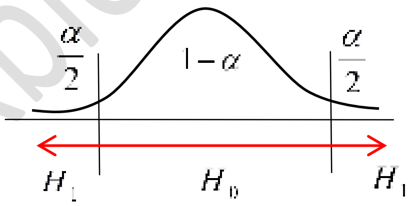
آزمون مقایسه میانگین دو جامعه آماری

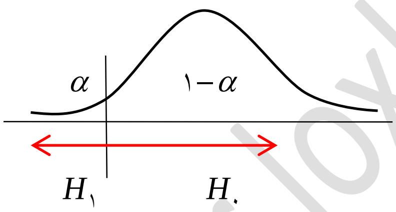
در این بخش سه حالت زیر بوجود می آید

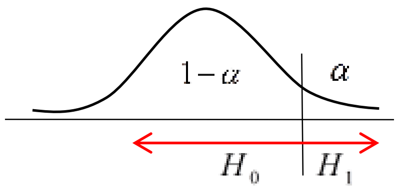
$H_0 \rightarrow \mu_1 = \mu_2$
$H_1 \rightarrow \mu_1 \neq \mu_2$

$H_0 \rightarrow \mu_1 \geq \mu_2$
$H_1 \rightarrow \mu_1 < \mu_2$

$H_0 \rightarrow \mu_1 \leq \mu_2$
$H_1 \rightarrow \mu_2 > \mu_2$







دو جامعه نرمال و انحراف معیار معلوم

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

دو جامعه نرمال و انحراف معیار نامعلوم $\sigma_1^2 = \sigma_2^2$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$SP\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

دو جامعه نرمال و انحراف معیار نامعلوم و نامساوی $\sigma_1^2 \neq \sigma_2^2$

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$d.f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

$n>30$ دو جامعه نرمال و انحراف معیار نامعلوم

$$z=\frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

$$d.f=n_1+n_2-2>30$$

مثال :

تحقیقی برای مقایسه روش آموزش متمرکز مدیران با روش غیر متمرکز آنان در دست برنامه ریزی است فرضیه عبارتست از : **(روش آموزش متمرکز برای مدیران بهتر از روش غیر متمرکز است)** ملاک سنجش فرضیه فوق مقایسه میانگین نمره های مدیرانی است که به روش متمرکز آموزش دیده اند نتایج نمونه گیری بشرح جدول زیر است ، بررسیها نشان میدهد توزیع نمره ها در هر دسته نرمال است در سطح خطای ۱٪ صحت فرضیه ها را آزمون کنید فرض تساوی واریانسها نیز برقرار است .

روش متمرکز	روش غیر متمرکز
$n_1=10$ $\bar{x}_1=52$ $s_1=12$	$n_2=15$ $\bar{x}_2=45$ $s_2=8$

۱- فرضها

$H_1 \to \mu_1>\mu_2$ ادعا

$H_0 \to \mu_1\le\mu_2$ نقیض ادعا

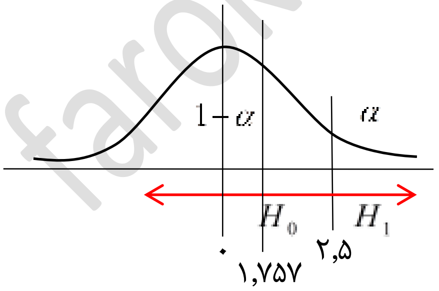
۲- آماره آزمون

چون تعداد نمونه کمتر از ۳۰ عدد است از t استفاده میکنیم .

$$t=\frac{(52-45)-0}{sp\sqrt{\frac{1}{10}+\frac{1}{15}}}=\frac{7}{9.762\sqrt{\frac{1}{10}+\frac{1}{15}}}=1.757$$

$$sp=\frac{\sqrt{s_1^2\left(n_1-1\right)+s_2^2\left(n_2-1\right)}}{n_1+n_2-2}\Rightarrow sp=\frac{(12)^2(10-1)+(8)^2(15-1)}{10+15-2}=9.762$$

۳- مقدار بحرانی



$$t_{\alpha},d.f$$

$$t_{0.01,23}=\pm2.5$$

H_1 رد میشود و H_0 تائید میشود .

آماره آزمون (۱٫۷۵۷) در مقایسه با مقدار بحرانی (۲٫۵) کوچکتر بوده و در ناحیه H_0 قرار میگرد بنابراین ادعا رد شده و نقیض ادعا پذیرفته میشود .

آزمون مقایسه زوجها

این آزمون برای ارزشیابی پیش آزمون و پس آزمون بکار میرود در این آزمون آزمودنیهای ما (منشاء پاخسگویی به نظرات درباره دو جامعه یکسان است) بنابراین با تشکیل زوجهایی از داده های گرفته شده از هر آزمودنی آزمون را اجرا مینمائیم در این آزمون داریم :

$$d_i = y_i - x_i$$

آماره آزمون =T

$$t = \frac{\bar{d} - 0}{s_{\bar{d}}}$$

$\bar{d}$=میانگین d_i

$s_{\bar{d}}$= انحراف معیار

مثال :

فرضیه ای عبارتست از (جو سازمانی موجود با جو سازمانی در وضعیت مطلوب اختلاف نامناسبی دارد) برای بررسی این فرضیه از ۵ مدیر نظر سنجی شده تا وضعیت موجود و مطلوب سازمان را ارزش گذاری کند در سطح خطای ۱٪ فرضیه را بررسی کنید ، نتایج حاصله بشرح زیر است :

مدیر	۱	۲	۳	۴	۵
x_i نمره وضعیت مطلوب	۵۰	۵۹	۵۰	۵۸	۵۰
y_i نمره وضعیت مطلوب	۴۰	۵۷	۴۷	۵۰	۴۸

فرض نرمال بودن برای نمره های جو سازمانی در دو وضعیت موجود و مطلوب برقرار است

۱- فرضها

$$H_0 \to \mu_d \geq 0$$
$$H_1 \to \mu_d < 0$$

۲- آماره آزمون

مدیر	۱	۲	۳	۴	۵	
x_i	۵۰	۵۹	۵۰	۵۸	۵۰	
y_i	۴۰	۵۷	۴۷	۵۰	۴۸	
d_i	-۱۰	-۲	-۳	-۸	-۲	$\sum d_i = -25$
$(d_i - \bar{d})^2$	۲۵	۹	۴	۹	۹	$\sum (d_i - \bar{d})^2 = 56$

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-25}{5} = -5$$

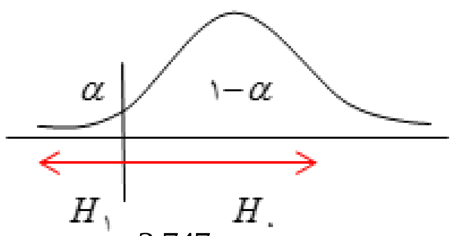
$$S_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1} = \frac{56}{5-1} = 14 \Rightarrow s_d = 3.742$$

$$t = \frac{-5}{\frac{3.742}{\sqrt{5}}} = -2.988$$

۳- مقدار بحرانی : آزمون از نوع دنباله چپ است که با توجه به $\alpha = 0.01$ ، منحنی آن بصورت زیر است .

$$\begin{cases} \alpha = 0.01 \\ d.f = n-1 \Rightarrow 5-1 = 4 \end{cases}$$

$$t_{0.01,4} = -3.747$$



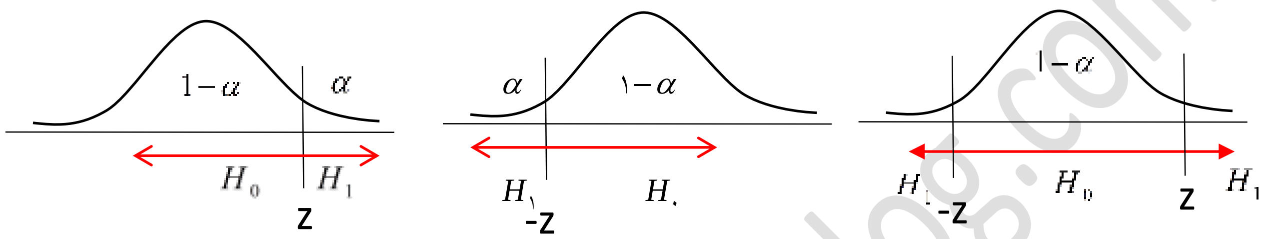
۴- تصمیم گیری :

آماره آزمون $t=-3.747$ در مقایسه با مقدار بحرانی در ناحیه H_0 قرار میگیرد بنابراین فرض H_0 در سطح خطای ۱٪ پذیرفته میشود بعبارت دیگر فرضیه پژوهشی رد شده ، نقیض آن تائید میگردد .

آزمون فرض نسبت موفقیت در جامعه (p)

در این آزمون نیز سه حالت متصور است .

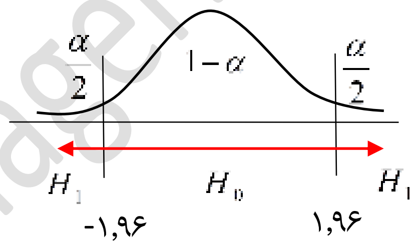
$$H_0 \rightarrow p = p_0 \qquad H_0 \rightarrow p \geq p_0 \qquad H_0 \rightarrow p \leq p_0$$
$$H_1 \rightarrow p \neq p_0 \qquad H_1 \rightarrow p < p_0 \qquad H_1 \rightarrow p > p_0$$



مثال :

فرضیه ای بدین صورت تدوین یافته ، ۶۰٪ مدیران شرکت از روش S۱ برخوردارند برای بررسی این فرضیه نمونه ۲۰۰ تایی از بین مدیران کشور انتخاب شده که نیمی (۱۰۰) از آنها از شیوه S۱ برخوردارند در سطح خطای ۵٪ درصد فرضیه را آزمون کنید .

$$H_0 \rightarrow p = 0.60$$
$$H_1 \rightarrow p \neq 0.60$$
$$\bar{p} = \frac{x}{n} = \frac{100}{200} = 0.5$$
$$z = \frac{0.5-0.6}{\sqrt{\frac{0.6(0.4)}{200}}} = -2.89$$
$$z_{\frac{\alpha}{2}} = z_{0.025} = \pm 1.96$$



H_0 تائید میشود چون کمتر از ۰.۶۰ است .

آزمون فرض نسبت موفقیت در دو جامعه آماری

در این حالت آماره آزمون عبارتست از

$$z = \frac{\left(\bar{p}_1 - \bar{p}_2\right) - \left(p_1 - p_2\right)}{\sqrt{\frac{\bar{p}_1\left(1-\bar{p}_1\right)}{n_1} + \frac{\bar{p}_2\left(1-\bar{p}_2\right)}{n_2}}}$$

مثال :

فرضیه ای بدین صورت تدوین یافته انگیزه توفیق طلبی بین دانشجویان شهرستانی و تهرانی به یک میزان است در سطح خطای ۱٪ فرضیه فوق را که داده های زیر برای آن جمع آوری شده آزمون کنید .

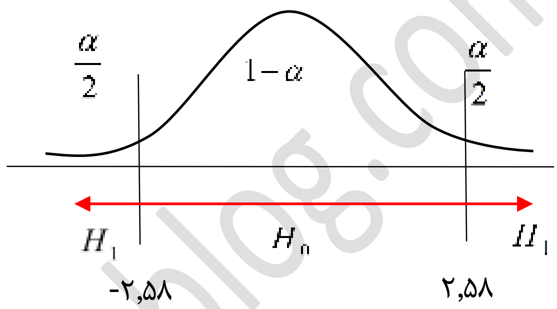
میزان توفیق طلبی / دانشجویان	بالا	پائین	مجموع
شهرستانی	۲۰۰	۱۰۰	۳۰۰
تهرانی	۲۵۰	۱۵۰	۴۰۰

$$H_0 \to p_1 = p_2$$

$$H_1 \to p_1 \neq p_2$$

$$\bar{p}_1 = \frac{x_1}{n_1} = \frac{200}{300} = 0.667$$

$$\bar{p}_2 = \frac{x_2}{n_2} = \frac{250}{400} = 0.625$$



$$z = \frac{(0.667 - 0.625)}{\sqrt{\frac{0.667(0.333)}{300} + \frac{0.625(0.375)}{400}}} = 1.154$$

$$z_{\frac{\alpha}{2}} = z_{0.005} = \pm 2.58$$

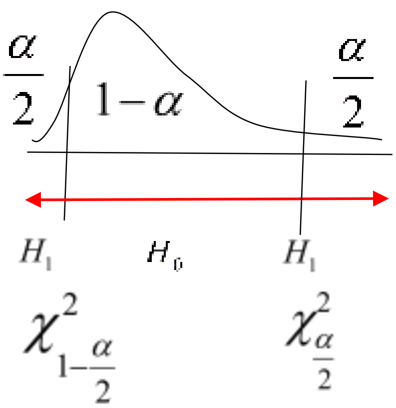
نتیجه گیری : مقدار آماره آزمون (z=۱,۱۵۴) در ناحیه H_0 قرار میگرد ، بنابراین فرضیه پژوهشی در سطح خطای ۱٪تائید میشود .

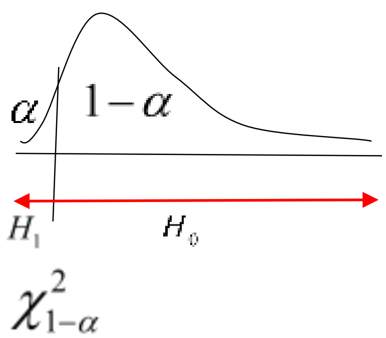
آزمون فرض آماری برای واریانس جامعه :

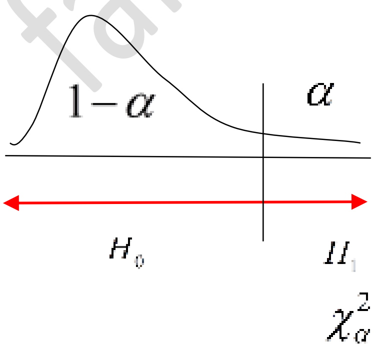
در این آزمون سه حالت زیر قابل اجراست :

$$H_0 \to \sigma_x^2 = \sigma_0^2 \qquad H_0 \to \sigma_x^2 \geq \sigma_0^2 \qquad H_0 \to \sigma_x^2 \leq \sigma_0^2$$

$$H_1 \to \sigma_x^2 \neq \sigma_0^2 \qquad H_1 \to \sigma_x^2 < \sigma_0^2 \qquad H_1 \to \sigma_x^2 > \sigma_0^2$$







$$\chi^x = \frac{(n-1)S_x^2}{\sigma_0^2}$$

مثال :

مدیر عامل بورس تهران ادعا کرده است ریسک (انحراف معیار) بازده سهام شرکتهای بورسی کمتر از ۵ تومان است برای آزمون این فرضیه ۲۵ شرکت بطور تصادفی انتخاب شده که میانگین بازده آن ۱۴ و انحراف معیار ۴ تومان است بازده شرکتهای بورس ، توزیع نرمال دارد فرضیه فوق را در سطح خطای ۵٪ آزمون کنید .

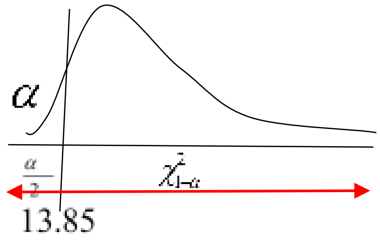
$\sigma_x = 5$

$H_0 \to \sigma_x^2 \geq 25$

$H_1 \to \sigma_x^2 < 25$

$$\chi^2 = \frac{(25-1)(4)^2}{(5)^2} = 15.36$$

$\chi^2_{0.95,24} = 13.85$



تصمیم گیری : آماره آزمون ($\chi^2 = 15.36$) در مقایسه با مقدار بحرانی ($\chi^2 = 13.85$) در ناحیه H_0 قرار میگرد ، بنابراین با ۹۵٪ اطمینان میتوان گفت فرض صفر پذیرفته میشود بعبارت دیگر فرضیه پژوهشی رد و نقیض آن پذیرفته میشود .

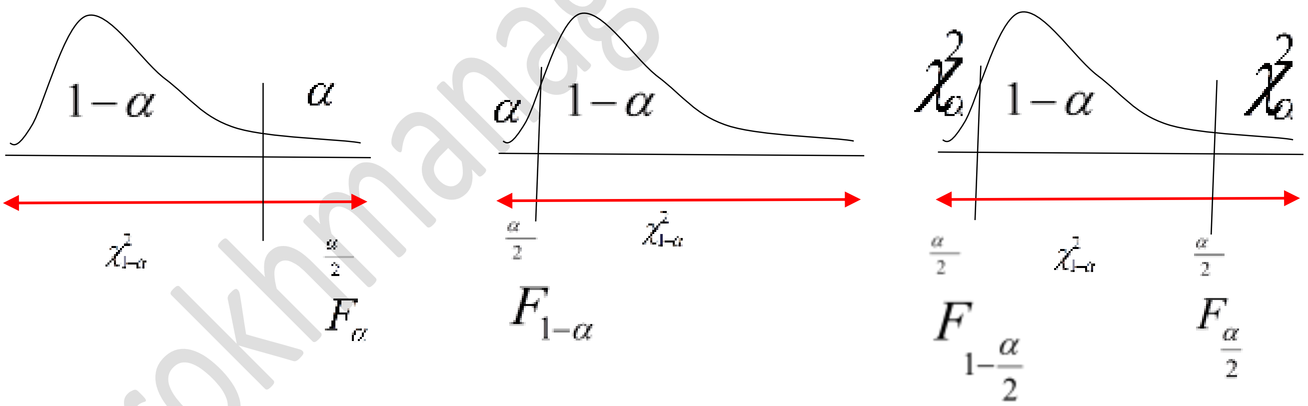
آزمون فرض آماری برای واریانس دو جامعه :

در این آزمون سه حالت زیر قابل اجراست :

$H_0 \to \sigma_1^2 = \sigma_2^2$
$H_1 \to \sigma_1^2 \neq \sigma_2^2$

$H_0 \to \sigma_1^2 \geq \sigma_2^2$
$H_1 \to \sigma_1^2 < \sigma_2^2$

$H_0 \to \sigma_1^2 \leq \sigma_2^2$
$H_1 \to \sigma_1^2 > \sigma_2^2$



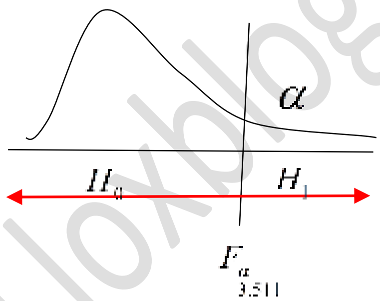
$$F = \frac{S_1^2}{S_2^2}$$

مثال :

فرضیه ای بدین صورت تدوین یافته ، پراکندگی وزن محصولات تولید شده بوسیله ماشین الف بیش از ماشین ب است برای آزمون این فرضیه داده هایی بشرح جدول زیر جمع آوری شده ، با فرض نرمال بودن تولیدات ماشین الف و ب فرضیه فوق را در سطح خطای ۵٪ آزمون کنید .

ماشین ب	ماشین الف
$n_b = 8$ $\bar{x}_b = 15.5kg$ $s_b = 2.25kg$	$n_a = 16$ $\bar{x}_a = 15kg$ $s_a = 4.5kg$

$$H_1 \rightarrow \sigma_a^2 > \sigma_b^2$$
$$H_0 \rightarrow \sigma_a^2 \leq \sigma_b^2$$
$$F = \frac{s_a^2}{s_b^2} = \frac{(4.5)^2}{(2.25)^2} = 4$$
$$F_{\alpha, d.f_1, d.f_2} \Rightarrow F_{0.05,15,7} = 3.511$$



تصمیم گیری : مقدار آماره آزمون در ناحیه H_0 قرار میگرد ، یعنی اینکه داده ها دلالت کافی بر تائید H_0 در سطح خطای ۵٪ ندارد بعلاوه چون H_1 بین کننده این ادعاست ، پس میتوان گفت ادعای مامور کنترل کیفیت در سطح اطمینان ۹۵٪ صحت دارد .

تحلیل واریانس

در تحقیقاتی که بدنبال بررسی یک ویژگی در چند جامعه آماری هستیم از روش تحلیل واریانس استفاده میکنیم روش تحلیل واریانس مبتنی بر مجموع مجذورات انحرافات کل داده ها از میانگین میباشد ، تحلیل واریانس قادر است این انحرافات را به احرافات ناشی از عامل مورد نظر و انحرافات ناشی از عوامل ناشناخته و خارج از تحقیق تقسیم نماید .

مجموع توان دوم انحراف معیار نسبت به میانگین

$$\sum_{j=1}^{n}\sum_{i=1}^{k}\left(x_{ij}-\bar{x}_{\cdots}\right)^{2}$$

مجموع توانهای دوم باقیمانده یا خطا	مجموع توانهای دوم ناشی از منبع ۳	مجموع توانهای دوم ناشی از منبع ۲	مجموع توانهای دوم ناشی از منبع ۱

تحلیل واریانس یک عامله

فرض کنید تعداد ضایعات ۳ ماسین را با هم مقایسه کنیم ضایعات هر کدام بشرح جدول زیر است .

ماشین اول	۸۶	۷۹	۸۱	۷۰	۸۴
ماشین دوم	۸۹	۸۲	۸۸	۷۶	۹۰
ماشین سوم	۸۲	۶۸	۷۳	۷۱	۸۱

میخواهیم بدانیم آیا تفاوت معنی داری بین کارکرد ۳ ماشین وجود دارد یا اینکه اختلاف بین آنها معلول تصادف است .

نکته :

درحالی کلی نمونه تصادفی به اندازه n از k جامعه آماری گرفته میشود فرض میکنیم متغییرهای تصادفی متناظر x_{ij} مستقل هستند توزیع های نرمال با میانگین μ_i و واریانس مشترک σ^2 دارد بنابراین هر مشاهده را میتوان $x_{ij}=\mu_i+e_{ij}$ نشان داد.

(e_{ij} متغییر تصادفی دارای توزیع نرمال با میانگین صفر و انحراف معیارمشترک σ^2)

مشاهدات را میتوان بصورت جامع تر نشان داد

$$x_{ij}=\underbrace{\mu}_{\text{میانگین}}+\underbrace{\alpha_i}_{\text{اثر تیماری}}+e_{ij}$$

$$\sum_{i=1}^{k}\alpha_i=0$$

در تحلیل واریانس فرض صفر با فرض مقابل زیر تعریف میشود :

$$H_0 \begin{cases} \mu_1 = \mu_2 = \mu_3 \ldots\ldots\mu_k \\ \alpha_1 = \alpha_2 = \alpha_3 \ldots\ldots\alpha_k = 0 \end{cases}$$

H_1 {
حداقل دو میانگین با هم برابر نیستند
حداقل یک α_i مخالف صفر هستند

$$SST = SS(\mathrm{Tr}) + \mathrm{SSE}$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_{..}\right)^2 = n\sum_{i=1}^{k}\left(\bar{x}_{i0} - \underset{\text{میانگین کل}}{\bar{x}_{00}}\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \overline{x_{i0}}\right)^2$$

با توجه به اینکه درجه آزادی تیمارها $k-1$ میباشد و همچنین درجه آزادی خطاها $k(n-1)$ میباشد بنابراین میانگین ها بصورت زیر محاسبه میشود .

$$MS(\mathrm{Tr}) = \frac{SS(Tr)}{k-1}$$

$$MSE = \frac{SSE}{k(n-1)}$$

آماره آزمون بصورت زیر از نسبت میانگین ها حاصل میشود .

$$F = \frac{MS(Tr)}{MSE}$$

میانیگن توان دوم تیمارها
میانیگن توان دوم خطاها

که بدین صورت محاسبه میشود :

$$F = \frac{MS(Tr)}{MSE} = \frac{SS(Tr)/(k-1)}{SSE/k(n-1)}$$

جدول تحلیل واریانس یک عامله برای مقایسه k تیمار

منبع تغییرات	مجموع توانهای دوم	درجه آزادی	میانگین توانهای دوم	F
تیمارها	$SS(Tr)$	$k-1$	$MS(\mathrm{Tr}) = \frac{SS(Tr)}{k-1}$	$F = \frac{MS(Tr)}{MSE}$
خطا	SSE	$k(n-1)$	$MSE = \frac{SSE}{k(n-1)}$	
جمع	SST	$kn-1$		

$$SST=\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}^2-\frac{1}{kn}.T..^2$$

$$SS(Tr)=\frac{1}{n}.\sum_{i=1}^{k}T_{i.}^2-\frac{1}{kn}.T..^2$$

$$SSE=SST-SS(Tr)$$

در سطح خطای $\alpha=0.05$ آزمون کنید که آیا اختلاف بین میانگین های تعداد ضایعات ماشینها معنی دار است یا نه .

۱- فرضها :

$$H_0\begin{cases}\mu_1=\mu_2=\mu_3........\mu_k\\ \alpha_1=\alpha_2=\alpha_3........\alpha_k=0\end{cases}$$

H_1 حداقل دو میانگین با هم برابر نیستند
حداقل یک α_i مخالف صفر هستند

۲- آماره آزمون :

ماشینها (تیمارها)	تعداد ضایعات					جمع ضایعات (T_i)	میانگین
ماشین اول	۸۶	۷۹	۸۱	۷۰	۸۴	۴۰۰	۸۰
ماشین دوم	۸۹	۸۲	۸۸	۷۶	۹۰	۴۲۵	۸۵
ماشین سوم	۸۲	۶۸	۷۳	۷۱	۸۱	۳۷۵	۷۵
						۱۲۰۰	

$$T..=T_1+T_2+T_3$$

$$T..=400+425+375=1200$$

$$\sum_{i=1}^{3}\sum_{j=1}^{5}x_{ij}^2=86^2+79^2+81^2+...81^2$$

$$SST=\sum_{i=1}^{3}\sum_{j=1}^{5}x_{ij}^2-\frac{T^2..}{kn}\Rightarrow 96698-\frac{1}{3\times5}(1200)^2=698$$

$$SS(Tr)=\frac{1}{n}\sum_{i=1}^{3}T_{i.}^2-\frac{1}{kn}.T^2..\Rightarrow\frac{1}{5}(400^2+425^2+375^2)-\frac{1}{3\times5}(1200)^2=250$$

$$SSE=SST-SS(Tr)\Rightarrow 698-250=448$$

منبع تغییرات	مجموع توانهای دوم	درجه آزادی	میانگین توانهای دوم	F
تیمارها	۲۵۰	۳-۱=۲	$\frac{250}{2}=125$	$\frac{125}{37.33}=3.35$
خطا	۴۴۸	۳(۵-۱)=۱۲	$\frac{448}{12}=37.33$	
جمع	۶۹۸	(۳×۵)-۱=۱۴		

۳- مقدار بحرانی :

$$\alpha = 0.05, n = 5, k = 3$$

$$F_{\alpha, k-1, k(n-1)} \Rightarrow \mathrm{F}_{0.05,2,12} = 3.89$$

۴- تصمیم گیری :

چون $F = 3.35$ از $F_{0.05,2,12} = 3.89$ کمتر است ، فرض صفر (فرض تساوی میانگین های جامعه) در سطح معنی دار ۵٪ را نمیتوان رد کرد و اختلاف در میانگین ها شانسی و تصادفی است .